DOCUMENT RESUME

ED 430 019 TM 029 732

AUTHOR	Gersten, Russell; Baker, Scott
TITLE	Effective Instruction for English-Language Learners: A Multi-Vocal Approach toward Research Synthesis.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.; National Association of State Directors of Special Education, Alexandria, VA.
PUB DATE	1999-04-00
NOTE	71p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23, 1999).
CONTRACT	HO23950013
PUB TYPE	Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Educational Environment; Elementary Secondary Education; English (Second Language); Focus Groups; *Instructional Effectiveness; Limited English Speaking; Meta Analysis; Qualitative Research; *Second Language Learning; *Synthesis; Teachers; Teaching Methods

ABSTRACT

The current state of the knowledge base on the effectiveness of specific instructional practices for English-language learners was examined by conducting an exploratory meta-analysis on eight studies of instructional interventions and using qualitative synthesis to supplement this analysis. The first data source was the eight intervention studies located through electronic searches. The second data source consisted of studies of the learning environments of English-language learners that focused on analyzing and describing instructional practices. The third data source consisted of information gathered from 5 professional work groups, which differed from focus groups in that all 44 participants were professionals. Within the eight empirical studies, no clear pattern emerged regarding effective instructional practices for English-language learners. A synthesis of these studies and other data sources suggests that a good English language development program should include three components: (1) a focus on the development of fluency and proficiency in English; (2) a concern with more formal, grammatical aspects of English; and (3) an emphasis on learning new academic content. (Contains 4 figures, 6 tables, and 79 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Effective Instruction for English-language Learners:

A Multi-Vocal Approach Toward Research Synthesis

ED 430 019

Russell Gersten and Scott Baker

Eugene Research Institute

Russell Gersten

2A

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM029732

Paper presented at the annual meeting of the American Educational Research Association. Montreal, Canada, April 19-23, 1999.

BEST COPY AVAILABLE

Effective Instruction for English-language Learners:
A Multi-Vocal Research Synthesis

In decrying the high levels of passion and low levels of rational discourse on the subject of the education of English-language learners, Ygazuirre (1998) notes that a shift in emphasis towards instructional issues and variables is likely to significantly improve both the level and the quality of discourse on the topic. Scholars such as Goldenberg (1994, 1996) and Moll (1988) argue convincingly that research needs to go beyond which language is used to teach English-language learners and beyond which model of bilingual education is best, and move toward a delineation of instructional methods for how to teach successfully. Goldenberg (personal communication, October 8, 1994) noted, for example, that "The language of instruction debate has so dominated discussion of how to best to serve the needs of language minority children that other issues, which are at *least* equally important, have not been adequately addressed." In 1989, Figueroa, Fradd, and Correa decried the lack of "a substantive body of empirical data on actual, well-controlled interventions . . . that improve the academic abilities of students who are English-language learners" (p. 17). By and large, despite huge interest in the topic, this is still the case.

The recent report by the National Academy of Sciences (NAS) (August & Hakuta, 1997) on the knowledge base of effective education for English-language learners laments that little has been learned from large-scale program evaluation studies, which have focused primarily on issues of the language used for instruction and the optimal time for introducing English. These program evaluation studies are problematic because of significant, and in many cases inherent, methodological limitations such as non-comparable control groups, lack of pretest data, and poor understanding of how specific programs actually get implemented in classrooms. The report also critiques the methodology utilized in the "effective schools" studies (e.g., Lucas, Henze, & Donato, 1990; Tikunoff et al., 1991) and concludes they are seriously flawed as well.

The Academy recommended that more research be conducted that examines the effects of specific instructional practices on the academic learning outcomes of English-

language learners. These studies should be linked to more specific research questions that can help guide practice such as: "What methods work best to give English-language learners access to the academic and social opportunities that native English speakers have while they are learning English? [Are] effective teacher practices for students generally sufficient to help English-language learners succeed in school?" (p.193).

After synthesizing the descriptive research on cognitive operations and process used by English-language learners, Fitzgerald (1995) concluded that the principles of reading instruction derived from advances in cognitive psychology that are commonly recommended for general classroom use are likely to be effective for English-language learners. She generated several hypotheses concerning the adaptations necessary for English-language learners. They are likely to require particular care in wording of questions, the pace of lessons will vary, and strategies used to activate background knowledge will be different.

The purpose of this synthesis was to examine the current state of the knowledge base on the effectiveness of specific instructional practices for English-language learners. The guiding question was seemingly straightforward: What do we really know about effective teaching practices for English-language learners in the elementary and middle school grades? Although it appears straightforward, the issues raised by the National Academy of Sciences (August & Hakuta, 1997) and Fitzgerald (1995), among others, about investigating instruction for English-language learners, suggests that providing an adequate answer is more difficult than it seems.

In 1994, when we began this project we were able to locate only four instructional research studies with valid designs. By 1997, that number had doubled to eight. The small number of controlled experimental studies led to our decision to conduct both an exploratory meta-analysis using the methodologies outlined by Cooper and Hedges (1994) on the set of eight instructional intervention studies, and to supplement this with qualitative synthesis techniques.

Developing a methodology to synthesize the extant knowledge base on this topic was difficult because it is a highly fragmented body of knowledge and there are deep-rooted conceptual differences among scholars and researchers who investigate this topic. Consequently, we utilized the framework for research synthesis articulated by Ogawa and Malen (1991) in their seminal article on *multi-vocal research synthesis.*

What is a Multi-Vocal Synthesis?

Ogawa and Malen (1991) called for integrative syntheses of a professional knowledge base on topics for which there is a scant empirical data, such as site-based management. They introduced *multi-vocal* synthesis methodology and urged its use for topics "characterized by a preponderance of diverse writings and a paucity of systematic investigations . . ." (p. 265).

> This strategy enables researchers to conduct . . . (an) open-ended search for relevant information, identify the major patterns associated with the phenomenon of interest, develop or adopt constructs that embrace the patterns, articulate tentative hypotheses about the meanings of the constructs and their relations, and refine questions and/or suggest conceptual perspectives that might serve as fruitful guides for subsequent investigations. . . (Ogawa & Malen, 1991, p. 271).

Using multi-vocal synthesis methods, researchers evaluate the methods and results of a given set of studies and use rigorous qualitative procedures to analyze "the words . . . in these diverse writings" (Ogawa & Malen, 1991, p. 265) to determine potential underlying belief systems and biases.

The research literature on effective instructional practices for English-language learners is appropriate to multi-vocal synthesis techniques since there are a variety of serious perspectives and little data. Each serious perspective needs a "voice" and to have its validity as a source of evidence considered. As Ogawa and Malen (1991) forcefully argued, "the literatures for some of the most prominent topics in education . .

. are multi-vocal. They are characterized by an abundance of diverse documents . . . [that are often] profuse [and] disparate . . ." (p. 266). Failure to attend to issues raised in the full gamut of literature, and in our view, failure to seriously examine patterns that emerge from experimental research through meta-analytic techniques lead to a limited understanding of issues.

Two formal data sources were included as "voices" in our multi-vocal synthesis: (a) experimental (i.e., intervention) studies and (b) descriptive studies of instructional practices that utilized classroom observations techniques. We did not weigh the findings from highly subjective or interpretative research as heavily as we did research conducted with valid experimental designs and reliable measures or qualitative studies that seemed to provide a more dispassionate analysis of issues raised.

Unique to our multi-vocal synthesis was the use of professional work groups as a third data source. These work groups, which included practitioners and researchers, helped identify relevant and irrelevant concepts for the integrative synthesis. We consider the input from these end-users a high priority and believe their contribution helped us develop an informed sense of the propositions and practice issues considered important by the most knowledgeable groups in the field. We believe their participation strengthens the validity of the interpretations that emerged and provides an important linkage between practice and research.

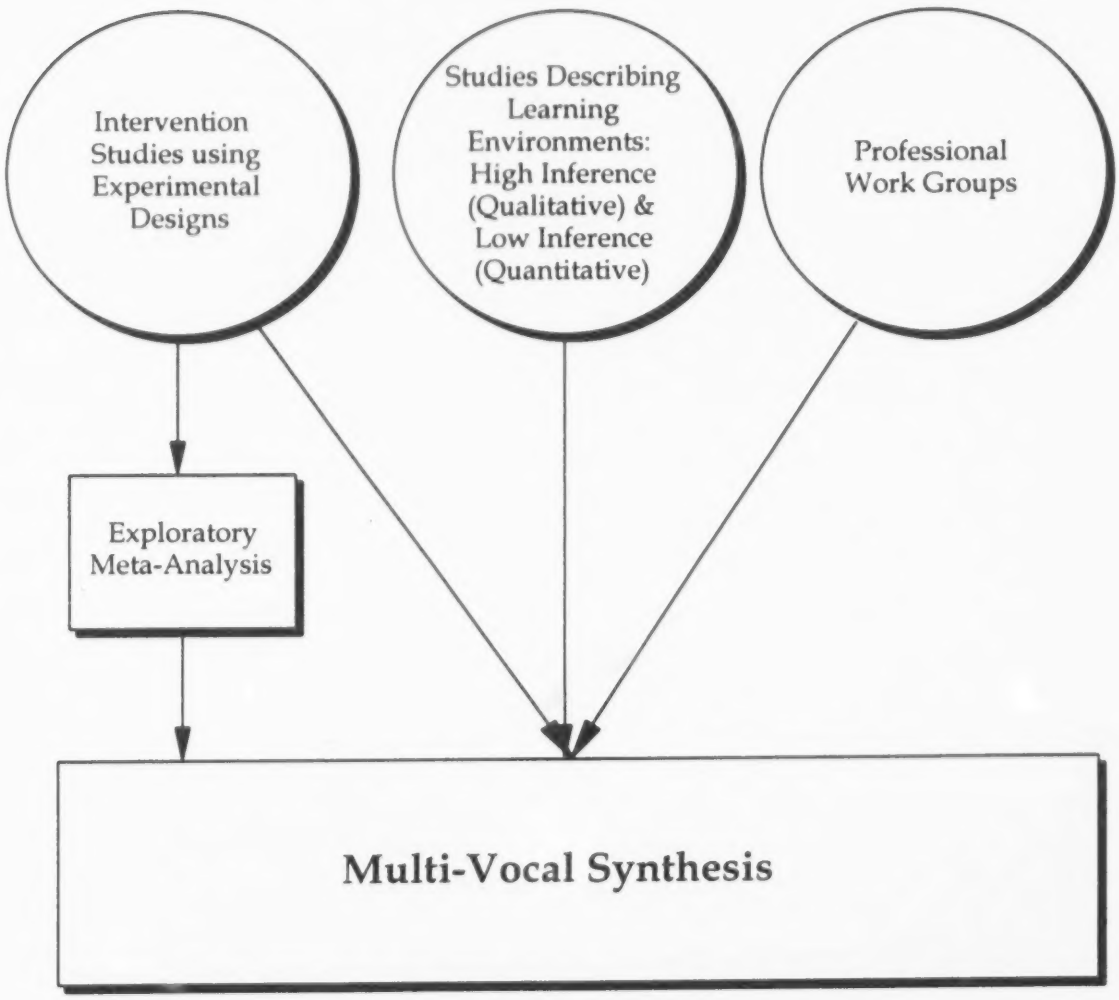
Data Sources

The various data sources and how they fit within the context of the integrated, multi-vocal synthesis are presented in Figure 1. As the figure shows, the first data source consisted of quantitative studies using experimental or quasi-experimental designs that examined the learning outcomes of specific instructionalapproaches. We analyzed these studies using traditional meta-analysis techniques (Cooper & Hedges, 1994).

The second and third data sources were qualitative in nature. The second source consisted of studies of the learning environments of English-language learners that focused on analyzing and describing instructional practices. Although some of the studies in this category used reliable

observation instruments to generate quantitative data, we relied on primarily qualitative methods to integrate our interpretation of these studies within the context of the overall syntheses.

Figure 1. Integrative Research Synthesis


Intervention Studies using Experimental Designs
Studies Describing Learning Environments: High Inference (Qualitative) & Low Inference (Quantitative)
Professional Work Groups
Exploratory Meta-Analysis
Multi-Vocal Synthesis

The third data source consisted of information gathered from the professional work groups. The studies we analyzed were important in our work with the professional work groups. Our research team read and coded these studies in preparation for the group discussions. In particular, we used our understanding of the studies to discuss the feasibility of applying specific principles of effective instruction in real classroom settings.

Our goal was to develop valid interpretations (Wolcott, 1994) from these disparate data sources. For the qualitative analyses, we followed Wolcott's dictum, "to open things up rather than seal them up . . . offering a new perspective gained after extended reflection. . . . The process can be stimulated and nurtured, but . . . it cannot be rushed" (Wolcott, 1994, p. 260). In fact, we devoted three years to this process.

Method

Literature Search

Studies were included in the synthesis if they focused on English-language learners in grades kindergarten through grade 8, and were conducted between 1985 and 1997.

For the intervention studies, we selected studies that used experimental and quasi-experimental designs that clearly measured the effect instructional variables had on students' academic outcomes. We utilized recognized standards (Wortman, 1994) for determining which studies were eligible for analysis. The intervention studies had to meet the following criteria:

1. Some objective measure of student performance was used to evaluate effectiveness of the intervention.
2. The study included a comparison group.
3. Sufficient data were reported for computation of an effect size (e.g., a study that reported posttest means but no standard deviations would not qualify).
4. If the study did not use random assignment of students to treatment, then pretest data must have been reported. In addition, pretest differences had to

be less than one-half of a standard deviation unit on relevant academic measures (so appropriates statistical adjustments could be made).

Studies that analyzed classroom learning environments were included in the analysis if they were based on classroom observations that used either reliable measures or a standard interpretive framework in their analysis. We divided studies in this category into those that relied on low-inference instruments for documenting and analyzing classroom practices and those that relied strictly on qualitative interpretations.

The following terms were used in our electronic searches for relevant studies: (a) English-language learner, (b) language-minority, (c) bilingual education, (d) limited-English proficiency, and (e) bilingual special education. The following electronic data bases were searched: ERIC, National Clearinghouse on Bilingual Education (NCBE), Dissertations Abstracts International, National Information Center for Children and Youth with Disabilities (NICHCY).

In addition, we utilized a range of Internet searches for publications using PLWeb (using the limiters *bilingual, instruction, and disability*). The following websites were searched for relevant studies: Urban Education Center for Research on Education, Diversity, and Excellence; Center for Research on the Education of Students Placed At Risk; and the Office of Bilingual Education and Minority Languages Affairs.

Manual scans of reference lists from selected research studies and other publications were also checked for additional studies. We also conducted a hand search of recent issues of 26 major relevant journals in education. Our final tally included 8 intervention studies[i] and 15 studies that analyzed classroom instruction.

Coding of intervention and descriptive studies. Our reading and coding of both the intervention and classroom observation and analysis studies proceeded in the following manner. One member of the research team read a study and entered the relevant data into a data display matrix (Miles & Huberman, 1994). This researcher reported key

features of the study to other members of the team. Team members then discussed the study's methodology, findings, and interpretations, and unresolved issues and questions were noted. The data display matrix was revised repeatedly during this series of initial interactions among research team members.

These discussions also provided the framework for how we would engage as a research team in the process of posing interpretations of the data, which in some instances were in sharp contrast with the interpretations posed by the study's author(s). Alternative interpretations and explanations of research findings are an essential aspect of qualitative research, one we felt should be central to our synthesis.

Professional Work Groups

We conducted five professional work groups across the U. S. Our major goal was to tap into participants' concepts about effective instructional practices for English-language learners. We reasoned that work group participants could identify what they saw as themes and problems in current practice, or problems with recommendations about best practice.

These professional work groups differed from focus groups in that: (a) all participants were professionals (teachers, staff development specialists, administrators, researchers) rather than consumers; and (b) our interactions with them were significantly longer (i.e., a total of 5 to 7 hours per group) than traditional focus groups. These participants and their professional positions are presented in Table 1.

Table 1. Educational Roles of Professional Work Groups

	Meeting Location				
Participant Roles	VA	CA	Wash., DC	FL	AZ
Researchers		3	6	1	3
Administrators	1				2
Teachers[a]	1	5	1	3	1
Psychologists	1			2	1
Staff Development	3		1	3	6
Total	6	8	8	9	13

[a]Teacher participants included bilingual, special, and general education teachers.

Although we initially proposed to conduct separate work groups with teachers and researchers, we decided that integrated heterogeneous groups were preferable. In preparatory meetings, teachers indicated they did not want to be "excluded" from the deliberations of researchers. These somewhat heterogeneous groups seemed to work well, and overall we felt they yielded more diverse discussions among participants. Our work group sessions were conducted in different geographic regions of the country because location was viewed as a variable that might influence how the groups responded.

We invited all researchers who had conducted research involving English-language learners supported by the Office of Special Education Programs. Several of these researchers had also conducted research supported by the Office of Educational Research and Improvement. School district personnel represented State Education Agency directors, school administrators, program administrators, and teachers.

Some of the work group sessions were audiotaped. In-depth notes by members of the research team were recorded at all sessions. In some cases, participants' written notes were also collected.

We asked each of the groups to respond to three broad topics:

1. *Four propositions* that helped guide the synthesis (see Figure 2). Each group provided feedback and suggestions for revising the propositions, including making major deletions or additions.
2. *Initial findings* from our literature search and analysis; and
3. A request for *real-world examples* illustrating a key principle or dilemma.

Using propositions to facilitate discussions and guide the synthesis. Our synthesis of the research began by developing a series of propositions about potentially useful instructional principles with English-language learners. We used these propositions, presented in Figure 2, to help guide our reading and analysis of the extent literature, and as a discussion catalyst for the professional work groups.

These propositions were derived in large part from the first authors' extensive observational research in classroom environments serving English-language learners (Gersten, 1996a; Gersten, 1996b; Gersten & Jiménez, 1994; Gersten & Woodward, 1994). This type of "grounding" is a cornerstone of qualitative research (Pressley, 1996; Strauss & Corbin, 1994).

The first three propositions most clearly focus on instructional strategies specifically with English-language learners. The fourth, cooperative learning, is a strategy that can be used effectively with all students. However, because of its potential to facilitate student discourse, we felt it warranted a specific investigative focus with English-language learners.

We sent the propositions to the participants prior to meeting with them. During the meeting we asked them to respond to, or comment on, each proposition. We reminded participants that these were *propositions,* and we expected and hoped they would change based on their feedback.

Beginning with the second work group, we added an additional task – the delineation of a list of principles and practices deemed to be productive for English-language learners. We presented the list generated by Group 2 (California) to Group 3 (Miami) as a work-in-progress so that each group could continue to refine it. Figure 3 outlines the process. The specific focus in each work group varied somewhat, primarily due to the unique composition of the members. For example, the San Diego work group, composed mostly of teachers and teacher supervisors, developed a rather practical set of practices. In contrast, the Washington, D.C. work group was primarily composed of researchers and was more theoretical.

In facilitating these meetings, we tried to "probe beneath the surface" (Blaunstein, 1995; Vaughn, Schumm, & Sinagub, 1996), to seek areas of discomfort with current practice and current theories advocated by state agencies and national organizations (such as NABE and CEC). We wanted the groups to articulate what they

saw as problems in current practice, and to provide the details of how teachers worked out solutions to vexing problems. To a large extent, this succeeded.

It is important to note that the professional work groups were conducted concurrently with our analysis of the published literature and issues raised in the literature helped frame questions we posed to the groups. Similarly, perspectives gained from the professional work group discussions guided and shaped our interpretations of issues and themes raised in the published literature.

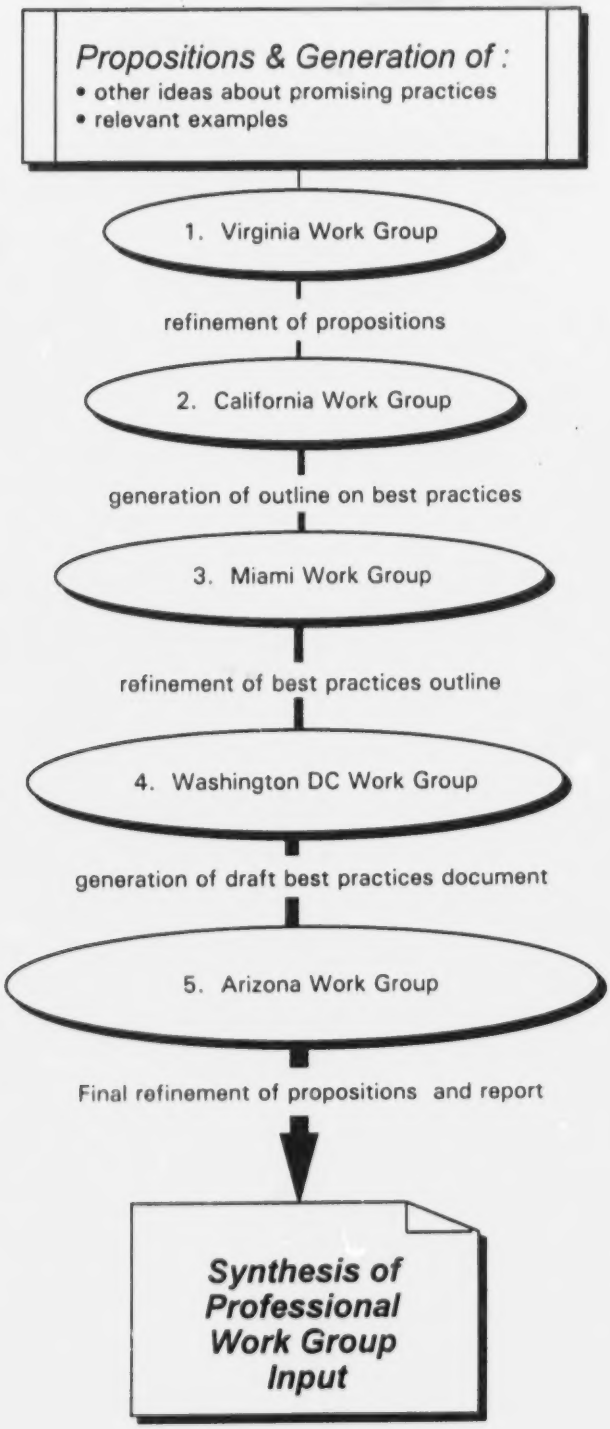
Figure 2. Propositions that Helped Guide the Synthesis and Promoted Discussion in the Professional Work Groups

1. Merging English Development (ESL or ESOL) Instruction with Content Instruction This proposition addressed the practice of merging English-as-a-second-language (ESL) instruction with content area instruction to develop students' knowledge of the more abstract language used in academic learning as opposed to conversational English. We proposed that some content areas (e.g., math, science) and some techniques are more promising than others.
2. Modulation of Cognitive and Language Demands This proposition was that effective teachers intentionally balance cognitive demands when the goal is to encourage English language expression (be it written or oral); in contrast, when the cognitive task is inherently demanding (e.g., a new science concept or complex literary content, such as character clues), teachers allow students to use their native language.
3. Transfer of Native Language Skills to English One very important issue for effective instructional practice with English-language learners is the issue of transfer, or applying native language skills to assist in learning a second language. This proposition was that explicit strategy instruction is required on how to access native language abilities and skills when learning content in English.

4. Structures that Support Cooperative Learning

This proposition was that there are certain *specific techniques* in implementing cooperative learning that lead to superior student outcomes.

Figure 3. Professional Work Group Process


Propositions & Generation of :
• other ideas about promising practices
• relevant examples
1. Virginia Work Group
refinement of propositions
2. California Work Group
generation of outline on best practices
3. Miami Work Group
refinement of best practices outline
4. Washington DC Work Group
generation of draft best practices document
5. Arizona Work Group
Final refinement of propositions and report
Synthesis of Professional Work Group Input

Overview of the Data Analysis Procedures for the Multi-Vocal Synthesis

A unique feature of the synthesis is the comparisons conducted within a data source (e.g., intervention studies), and across data sources (i.e., intervention studies, descriptive observation studies, and the professional work groups). The major techniques we used to generate and refine themes and issues, and to develop valid interpretations from the data sources are based on standard principles of qualitative research.

For a period of 18 months, we used an iterative process of forming tentative interpretations, re-reading and re-examining study features, posing new interpretations, looking for corroboration in other data sources, and often, returning to the original study. In this process, we found the use of memos and mini-reports on critical articles useful. In Appendix 1 we present a list of other sources read that helped us understand and contextualize specific findings and data patterns.

We borrowed freely from the suggestions of major qualitative methodologists such as Wolcott (1994) and Miles and Huberman (1994). In particular, we followed guidelines for integrative syntheses using qualitative methods (Noblit & Hare, 1988; Ogawa & Malen, 1991). We also learned from a few qualitative integrative reviews of aspects of the published literature (Fitzgerald, 1995).

Each study's features were entered into a data display matrix that reflected in-depth analysis by the whole group. It included not only the surface features of the study (e.g., subject area, language of instruction, grade level, length of treatment, number of observations), but also our appraisal of the validity of the assertions made by the authors based on the design of the study and measures, and our sense of the major themes or issues or findings that emerged. The data display matrices were electronic and periodically updated based on subsequent review by research team members following procedures suggested by Miles and Huberman (1994).

Studies were clustered and re-clustered electronically according to the questions that began to emerge. Occasionally, we added categories as themes emerged and changed. For

example, we added the category *Language Used During Instruction by Teacher and by Students* after the first wave of coding. The coding process and use of data display matrices allowed us to examine the features of individual studies in relation to the entire body of data sources.

We intentionally did not stick to one clustering of studies at a time, but rather "roamed" through the data set to explore trends, issues, and hypotheses following practices recommended by Noblit and Hare (1988). As mentioned previously, this helped us juxtapose disparate data sources to test our interpretations.

As a research team, we constantly revisited our set of interpretations and inferences, and did additional reading and re-reading to explore alternatives. We used the following set of questions to guide the development and prioritization of our interpretations:

1. Which interpretations and recommendations have the most promise, based on level of evidence and strength of research support?
2. What recurring issues pervade the set of studies and other data sources?
3. What are the most frequent conclusions/findings/interpretations (i.e., areas of convergence)?
4. What are areas for further inquiry/research efforts, as identified by researchers? As identified by experts in the field?

As we began to note patterns within and among the data sources, we started looking for what Miles and Huberman (1994) refer to as the "critical case" (i.e., the case that "proves" or solidifies a finding or interpretation) as well as "potential disconfirming cases." Serious analysis of potentially disconfirming cases can actually "teach us much about the assumptions that guide various studies" (Noblit & Hare, 1988, p. 62).

Another type of integration involved studies with findings that appeared to "refute" one another. Similarly, we noted several studies where the descriptions seemed quite rich and valid, but our interpretations diverged considerably from the authors' (e.g., Perez, 1994; Ruiz, 1995). These conflicts led to in-depth explorations of alternative hypotheses, as recommended by qualitative methodologists (e.g Miles & Huberman, 1994).

We tried to understand how and why researchers may have viewed things differently. Noblit and Hare (1988) suggest that after examining and noting differences in interpretations, researchers may come to recognize the "descriptive account of the other as reasonable" (Noblit & Hare, 1988, p. 54). We found this aspect of analysis and interpretation to be among the most difficult, yet rewarding, tasks of the synthesis.

Through this process, a study's features and characteristics often took on different weight than they did in earlier discussions. Sometimes, a new variable or construct emerged as important. Noblit and Hare (1988) discuss this as a process in which "what was hidden becomes apparent; we better understand what was studied by making clinical inferences from the studies" (p. 75).

For example, after critical examination of the first set of five studies and summarization of the work group findings, the following emerging issues continued to arise in our weekly discussions:

1. General student engagement versus student intellectual engagement.
2. Whether student "talk" (or discourse) was in a student's first or second language, and the value authors placed on each (i.e. some authors clearly favored native language discourse without providing a clear rationale).
3. Problems defining the presence of cross-culturally competent or culturally relevant teaching.

We followed Noblit and Hare's (1988) dictum that each study read and analyzed helps understand and interpret the next study, as well as to re-analyze what was previously read and discussed. They called this process "*reciprocal translation.*"

Initial analysis of the professional work group data involved writing up each of the sessions and compiling all the work group data into one draft summary. The data were sorted using a software program, Hyperqual, and across the work groups, data chunks were placed into six general categories; Instructional Strategies, Collaboration, Supports, Culture, Other Unresolved Issues, and Ideas for Dissemination and Communication. After this initial

"chunking" of the data, key patterns for each of the categories were examined in depth to get a sense of areas of agreement and convergence on effective instructional practices among researchers, teachers, and administrators.

The ultimate goal—and the objective of the current level of analysis—was to use these data to confirm underlying themes from the literature. We used data from the groups to assess what was working and not working for teachers and their students and to specify areas for which curriculum seemed particularly weak. We listened for instances where teachers or researchers talked about past practices that were discontinued due to administrative fiat, but which there was still the belief these practices could be effective. Finally, we listened for areas where the groups seemed conflicted or confused.

Throughout this process, we continued to revise the documents we used to record our interpretations (e.g., the propositions, the data display matrices, the list of instructional approaches). In Figure 4, we present an example of how a working proposition was refined to lead to a more meaningful and potentially useful one over the two-year period.

To reiterate, the major principles used in the multi-vocal data analysis and interpretation were:

1. Significant input from practitioners for generation and refinement of interpretations (Ogawa & Malen, 1991).
2. *Triangulation* across various data sources (Patton, 1990).
3. Use of *Propositions* generated from immersion in environment and published research to help guide discussion and analysis.
4. *Constant-comparative* method of traversing data sources to develop and refine interpretations (Strauss & Corbin, 1994). Conscious juxtaposition of disparate studies (Gersten, Morvant, & Brengelman, 1995; Noblit & Hare, 1988).
5. Serious entertaining of rival hypotheses (Noblit & Hare, 1988).
6. Reciprocal translation (Noblit & Hare, 1988).

Figure 4: Evolution of a Working Proposition

The original proposition was: *It is effective to merge English as a second language instruction with content instruction to develop students' knowledge of the more abstract language used in academic learning as opposed to conversational English.*

This proposition was presented to each of the professional work groups. In no case did participants indicate that they knew more than the research team about which content area worked best. Thus, we feel, with some confidence that there is neither empirical evidence or even "craft knowledge" indicating which content area is best for merging ESOL instruction and content instruction, still holds.

However, in the conversations that ensued, other issues rapidly surfaced. In particular, the California professional work group engaged in a detailed discussion concerning how content area ESOL almost invariably fails to provide adequate time for language learning. In other words, participants felt that teachers often emphasize content acquisition over building English language abilities. This point was also emphasized by several researchers doing observational research in California schools (Echevarria & Graves, 1998).

We then revisited and discussed the cycle of English Language Development (ELD) methodologies used in American schools in the past 20 years and assessed their strengths and weaknesses. To summarize briefly: Problems with the traditional approach that stressed syntax and grammar was that generalization and transfer were often limited (Tharp & Gallimore, 1988). Problems with the "natural language" approach used in the 1980s were that conversations between an adult and 32 students were rarely natural, natural language didn't easily fit the conditions and constraints of classrooms, and cognitive demands were too low. The California group concluded that a balance of the three might well be ideal, rather than the current move towards only using content area ESOL as the sole means of second language acquisition.

This proposition was further refined in subsequent work groups. In one of these groups it was noted that this proposition linked nicely with the working Proposition 2 about intentional modification and modulation of cognitive demands depending on whether a teacher's goals were primarily cognitive in nature or language learning.

Results and Discussion

We present the analysis and findings for each of the data sources in the following sections. In the first two sections, we present the set of studies included in the meta-analysis and then the set of studies describing and analyzing instruction in classrooms. In the third section, we present our findings and interpretations from the professional work groups. Following these separate sections, we present the integration of the findings across the three data sources.

Data Analysis Procedures for Exploratory Meta-Analysis

The basic index of effect size used in this exploratory meta-analysis was Glass' delta, defined as the difference between the treatment and comparison group means divided by the comparison group standard deviation (Glass, McGaw, & Smith, 1981). For studies that reported pre- and posttest scores, we calculated posttest effect sizes adjusting for pretest performance using procedures suggested by Wortman and Bryant (1985). In this adjustment, the effect size is calculated in the following way: The quantity of the pretest experimental mean minus the pretest comparison mean is divided by the comparison pretest standard deviation. This quantity is subtracted from the unadjusted posttest experimental mean minus the unadjusted posttest comparison mean divided by the comparison posttest standard deviation.

For the four studies that reported posttest data only, we subtracted the mean of the comparison group's unadjusted posttest score from the mean of the treatment group's unadjusted posttest score and divided by the comparison group's posttest standard deviation. In the one study that did not report means and standard deviations (i.e., Henderson & Landesman, 1995), effect sizes were estimated from the *F* ratio following the procedures described in Rosenthal (1994). In Table 2, we present each of the 8 studies in the exploratory meta-analysis and how they were categorized for the purpose of the meta-analysis.

All the participants in the studies were English-language learners. In two studies, the participants also had learning disabilities. Six studies were conducted in elementary schools, two in middle schools. Five of the eight studies were conducted primarily in English, the students'

second language (marked with a double asterisk in the reference section.) Three studies were conducted primarily in Spanish, the students' native language (marked with a triple asterisk in the references).

Six of the eight studies used a between-subjects group design with one experimental condition and one comparison condition. One study included three intervention conditions with no traditional comparison group (Waxman, de Felix, Martinez, Knight, & Padrón, 1994). One study utilized a counterbalanced, within-subjects design (Echevarria, 1995).

All but one study (i.e., Cardelle-Elawar, 1990) utilized multiple dependent measures. With multiple outcome measures, a single average effect size was computed so that each study included only one effect size per aggregation. Cooper and Hedges (1994) underscore that this approach retains as much data as possible while minimizing any violations of the assumption that the data points used in the analyses are independent.

Findings of the Exploratory Meta-Analysis

In Table 3, we present effect size data for the 8 studies for each dependent variable. When possible, the effect sizes in this table are adjusted for pretest differences. In total, 19 effect sizes were calculated, ranging from -.56 to 1.95. The mean effect size was .31, the median, .25. Overall, the median effect size indicated the interventions had a positive, but small impact on student learning.

The effect sizes were aggregated across the 8 studies to consider important variables for analysis and interpretation. Because of the small number of effect sizes included in any particular meta-analytic comparison, we decided to base our interpretations on the median effect size in each analysis. The median typically is a more reflective measure of central tendencies in small sets of studies.

Table 2. Features of Intervention Studies

Study	Number of students	Grade	Duration of treatment	Assignment of Students to Treatment	Primary Language of Instruction	Content Area	Type of Dependent Measure	Language of Dependent Measure
Cardelle-Elawar, (1990)	80	6th	6 hours (over 3 weeks)	Intact groups	English	Math	Experimenter developed	English
Echevarria, (1995)	5	2nd-3rd	25 lessons (over 1 year)	Counter balanced	Spanish	Reading	Experimenter developed	Spanish
Goldenberg et al., (1992)	10	K	1 year	Random	Spanish	Reading	Experimenter developed	Spanish
Henderson & Landesman, (1995)	102	7th-8th	2 years	Intact groups	English	Math	Experimenter developed	English
Klingner & Vaughn, (1996)	26	7th-8th	27 days	Random	English	Reading	Experimenter developed *and* Standardized	English
Muniz-Swicegood, (1994)	95	3rd	6 weeks	Random	Spanish	Reading	Standardized	English and Spanish
Saunders et al., (1998)	36	5th	1 year	Intact groups	English	Reading, language arts and math	Standardized	English
Waxman et al., (1994)	219	1st-5th	6 months	Intact groups	English	Reading and language arts	Standardized	English

Table 3. Individual Effect Sizes Used in the Aggregation of Effect Sizes

Study and Description of Experimental Intervention	Mean Effect Size for All Measures	Effect Size for Individual Measures
Cardelle-Elawar, M. (1990). *Elementary School Journal, 91*(2), 165-175. Students were pretested on 20 math story problems and responded to 8 statements that reflected difficult aspects of the problem. Teachers taught to give feedback to students over 3 weeks that addressed pretest errors and student statements. Errors classified into difficulties in (a) knowledge of English; (b) understanding how to produce math statements from prose; (c) determining procedural sequence for solving problems and; (d) accurate computation.	1.95	Math story problems:1.95
Echevarria, J. (1995). *Exceptional Children, 61*, 536-552. Teachers taught using *Instructional Conversations* format. Focus on themes and expanded discussions, inferential comprehension, and elaborated student dialogue.	-.22	Oral Retell .25 Idea Units -.36 Literal Recall -.56
Goldenberg, C., Reese, L., & Gallimore, R. (1992). *American Journal of Education, 100*(4), 497-536. Teachers introduced a new *libro* (small book) every three weeks, reading story several times over a few days. Books sent home and parents told to enjoy them and not use them to teach decoding or word recognition. However, parents used the *libros* to teach decoding and accurate reading.	.79	Letters & Sounds .72 Other Lit. Measures .86
Henderson, R. W., & Landesman, E. M. (1995). *Journal of Educational Research, 88*(5), 290-300. Combination of explicit teaching of concepts and thematic applications. Themes included fine arts, the ocean, crime, careers, the environment, world issues, sports, and the future. Students worked in cooperative groups of 5, with assigned roles (e.g., manager, engineer, accountant). Each job had duties that required math skills. These aligned with scope and sequence in textbook.	.18	Math Concepts .22 Math Computation .14
Klingner, J. K. & Vaughn, S. (1996). *Elementary School Journal, 96*(3), 275-294. All students taught Reciprocal Teaching in groups of 6 or 7; students learned 6 reading comprehension strategies. Content and discussions primarily in English, but students encouraged to use Spanish to understand concepts. Students in peer tutoring group then tutored grade 6 students in reading comprehension strategies. Tutees eventually became the "teacher." Students in cooperative learning group worked in groups of 3 or 4. Used Reciprocal Teaching framework, with students serving as facilitators.	.12	Standardized Reading-.2 Reading Comp. .32 Reading Strategies .25
Muñiz-Swicegood, M. (1994). *Bilingual Research Journal, 18*(1-2), 83-97. Students read text (in Spanish) and were taught to generate questions at the end of each paragraph. Students lead dialogues in small groups, asking questions of one another. Group size reduced until students worked in pairs. Finally, students generated questions on their own and discussed questions and answers with teacher individually.	.31	English Reading .39 Spanish Reading .22
Saunders, W., O'Brien, G., Lennon, D., & McLean, J. (1998). In Gersten, R., & Jimenez, R. (Eds.), *Effective strategies for teaching language minority students*. Belmont, CA: Wadsworth. Three critical intervention components were included in literacy instruction: 1. Teaching background knowledge through vocabulary, relating concepts to personal experiences, and academic applications. 2. Reading and working with text, with a focus on building vocabulary, developing comprehension strategies, and writing about the content. 3. Developing a piece of writing that expresses story theme provided students with extended opportunities to write. Feedback from teachers and peers used to help students revise and improve written products.	.75	English Reading .47 English Language 1.03
Waxman, H. C., de Felix, J. W., Martinez, A., Knight, S. L., & Padrón, Y. *Bilingual Research Journal, 18*(3-4), 1-22. In *Effective Use of Time* condition (Stallings, 1980) small teacher groups learned to use time effectively and provide more explicit instruction. Instruction broken down into concrete behaviors with observations and coaching provided. Most interactions provided in large group training sessions. In *ESL in Content Area* condition (Chamot & O'Malley, 1989), the focus was on problem solving in science, math, and reading in English and Spanish. Some concepts explained in Spanish - this amount reduced over time. Teacher training sessions addressed	.04	Reading: .27 Language: -.20

cognitive strategies.		

In examining the types of dependent measures used, the effect sizes are small for both standardized measures (median effect size = .17) and experimenter developed measures (median effect size = .29). There is a good deal of variability, however, ranging, among experimenter developed measures, from -.22 for *Instructional Conversations* (Echevarria, 1995) to 1.95 for tailored feedback in math (Cardelle-Elawar, 1990). Our results did confirm what other meta-analyses (Rosenshine & Meister, 1994; Swanson, in press) have found, namely that experimenter-developed measures typically result in higher effect sizes than standardized tests. The magnitudes of the difference between the two was much less than is typically reported in other meta-analyses, however.

Effect sizes are also relatively low when comparing across content areas. The median effect size for reading studies was .29 while the median for language was .62 and math was 1.07. It is important to note that although the test manuals indicate the language subtests are measures of "Language," the focus is exclusively on written *conventions* of language (i.e., capitalization, use of commas, tense agreement, sense of what constitutes a paragraph) and they do not assess critical aspects of language learning, such as vocabulary and the correct use of syntax.

Withstanding these caveats, the pattern of findings (i.e., the effect size in math was higher than language, and language was higher than reading) is one that invariably is found in national and local assessments with English-language learners (Gersten & Woodward, 1995; Natriello, McDill, & Pallas, 1990). Latino students, for example, tend to perform very poorly on English-language reading tests, and much higher on mathematics tests.

Nature of the Intervention

Our data indicate that when the independent variable(s) in a study was well defined and/or implementation was carefully assessed or monitored, the median effect sizes were higher than they were when the nature of the instructional interventions

were more loosely defined (.31 versus .18). In three of the eight studies, students were randomly assigned to treatment conditions. When random assignment was used, the effect sizes were smaller than when quasi-experimental designs were used (.31 versus .47). In other words, by and large, studies that employed higher quality comparison groups tended to have somewhat weaker effects. This matches findings from Swanson (in press), who investigated instructional interventions for students with learning disabilities.

A Potential Pattern Related to Instructional Framework

One of the most interesting patterns of findings, although clearly tentative, occurred when comparing instructional approaches that were largely based on extensions and adaptations of effective teaching research of the 1980s (Brophy & Good, 1986) with approaches that attempted to integrate aspects of social constructivism and cognitive psychology. The former focused squarely on the development of reading, writing, and math problem solving, and, to a high degree, relied on teacher-led instruction, with modeling, immediate individualized feedback, and opportunities for extended practice. Approaches focusing on constructivism and cognitive psychology are characterized by active student participation in the development of solutions to academic problems, student dialogue and opportunities for expanded academic discourse, and a focus on the development of metacognitive strategies as a critical component in solving academic problems.

The three intervention studies with the largest overall effect sizes, and the only studies with effect sizes that Cohen (1988) would describe as moderate to large in magnitude (i.e., Cardelle-Elawar, 1990; Goldenberg et al., 1992; Saunders, O'Brien, Lennon, & McLean, 1998) used intervention components that were extensions of instructional approaches described by Brophy and Good (1986), Rosenshine (1986), and others in the 1980's. These studies extended these practices to be responsive to the unique learning needs of English-language learners. Saunders et al., for example,

focused heavily on teaching students vocabulary critical to understanding academic content and provided extensive feedback (both teacher and peer) to students on their use of English to communicate their understanding.

In Cardelle-Elawar (1990), the sophisticated use of the recurrent finding from the effective teaching literature, that strategy feedback invariably leads to higher levels of math achievement, was replicated (Good & Grouws, 1977; Hiebert & Wearne, 1993; Kelly, Gersten, & Carnine, 1990). In Cardelle-Elawar's application of this principle, strategic and extensive feedback was provided to teach students how to attack solving math story problems. These instructional strategies focused on accurate computation and comprehending subtle aspects of language that provided critical information in understanding the problem.

Goldenberg et al. (1992) designed an instructional approach for parents to use at home to help their child learn to read. As designed, the approach had a strong constructivist orientation. Parents were supposed to read small books with their child strictly for enjoyment. In fact, they were explicitly instructed *not* to use the books for teaching decoding or accurate word reading. The observations Goldenberg and his colleagues conducted in the homes, however, indicated that decoding and accurate word reading were precisely the components of instruction that parents focused on most.

The Saunders et al. (1998) instructional approach was actually an artful blending of approaches adapted from advances in cognitive psychology with approaches that extended instructional principles discussed in the effective teaching literature. Consistent and moderate to strong positive effects on standardized measures in both reading and language were found.

In contrast, the studies with the two lowest effect sizes (the *CALLA* approach in Waxman, et al., 1994; Echevarria, 1995), which were zero and negative in magnitude, included two of the most frequently recommended approaches for teaching English-

language learners, Cognitive Academic Language Learning Approach (*CALLA*) and *Instructional Conversations*.

Both of these instructional approaches devote extensive opportunities for students to engage in what are claimed to be challenging language develop activities, which are assumed to lead to superior academic achievement outcomes. Both of the approaches also resulted in patterns of effect sizes that defy simple interpretations. In the Waxman et al. (1994) study, Chamot and O'Malley's complex cognitive strategies, in *CALLA*, led to negative effects in reading (.27), but positive effects in language (.20). In our review of *CALLA* instructional materials, this makes sense. The program has excellent English Language Development activities, but content covered appears to be less than would typically occur.

The Echevarria (1995) study presents a complex if somewhat puzzling pattern of findings. The *Instructional Conversations* intervention was intended to increase students' language use during reading instruction, and it was hoped it would have a positive impact on comprehension, knowledge of story structure, and the ability to elaborate on stories read and discussed. Ironically, the effect size was small and positive for story structure (.25), but negative both for the richness of the narrative retellings (-.36) and for literal comprehension of the story (-.56). it is interesting that traditional basal reading instruction led to larger and moderately strong effect sizes for literal comprehension and the richness of story retellings.

From our reading of the study, it appears that Echevarria omitted, or significantly reduced the importance of, the explicit instruction component of *Instructional Conversations*. This is increasingly viewed as an important component of *Instructional Conversations*, as articulated by Saunders and Goldenberg (1996). In Echevarria's study, teachers may have been well trained to engage in the aspects of *Instructional Conversations* that focus on enriching student discussions of narrative text,

but in the absence of explicit instruction to help students comprehend critical information, their understanding and deep processing may have been restricted.

Single Subject Study

To analyze the single-subject study (i.e., Rousseau, Tam, & Ramnarain, 1993), we interpreted the percentage of non-overlapping data points (Mastropieri, Scruggs, Bakken, & Whedon, 1996) in each condition for two dependent measures: the percentage of words students read correctly, and the percentage of comprehension questions students answered correctly. We summarized the data across the five students who received each of the treatments in the alternating treatments design.

The unequivocal finding was that the Vocabulary condition was more effective than the Listening Preview condition in improving students' reading. Students performed much higher on both percent of words read correct (a mean of 95% vs. 50%) and questions answered correctly (97% vs. 68%). In the combined Vocabulary and Listening Preview condition students scored 100% on both dependent measures. At the individual student level, the Listening Preview condition seemed to be extremely ineffective for 2 of the 5 students, whereas the Vocabulary intervention was successful for every student. The results of this study support the previous conclusion about extensions of effective teaching showing effects superior to studies using more constructivist techniques.

Brief Overview of Research Describing Instructional Environments

Because our goal, in part, was to "provide a detailed portrait of the phenomenon (being studied) . . . the ways it appears to operate, and the patterns observed in natural settings" (Ogawa & Malen, 1991), it was important to include in the synthesis those studies that described and analyzed actual practices observed during instruction.

Fifteen studies were reviewed in this category. These studies focused on observations of students during instruction, and attempted to assess engagement levels, student academic and oral language interactions, and to describe the types of

instructional arrangements used to facilitate learning, such as independent seat work, and small or whole group instruction. Four studies in this category used low inference observation systems with reported reliability coefficients to assess aspects of instruction, student language use, and engagement (Arreaga-Mayer & Perdomo-Rivera, 1996; Padrón, 1994; Ramírez, 1992; Tikunoff et al., 1991). The 11 other studies used a wide array of qualitative or interpretive methodologies to better understand dimensions of practice.

Information and analyses gleaned from these 15 studies were used to develop themes for the multi-vocal synthesis, interpret issues raised by the professional work groups, and provide useful context for understanding aspects of the intervention studies. Table 4 references each of these studies.

Table 4. Studies Describing Instructional Environments

Low inference (N=4)

Arreaga-Mayer, C., & Perdomo-Rivera, C. (1996). Ecobehavioral analysis of instruction for at-risk language minority students. *Elementary School Journal, 96*(3), 245-258.

Padrón, Y. (1994). Comparing reading instruction in Hispanic/limited-English-proficient schools and other inner-city schools. *Bilingual Research Journal, 18*(1-2), 49-66.

Ramírez, D. J. (1992). Executive summary. *Bilingual Research Journal, 16*(1&2), 1-62.

Tikunoff, W., Ward, B., von Broekhuizen, D., Romero, M., Castaneda, L., Lucas, T., & Katz, A. (1991). *Final report: A descriptive study of significant features of exemplary special alternative instructional programs.* Los Alamitos, CA: The Southwest Regional Educational Laboratory.

High inference (N=11)

Echevarria, J., & McDonough, R. (1993). *Instructional Conversations in special education settings: Issues and accommodations.* National Center for Research on Cultural Diversity and Second Language Learning.

Gersten, R. (1996a). Literacy instruction for language-minority students: The transition years. *Elementary School Journal, 96*(3), 227-244.

Gersten, R., & Jimenez, R. T. (1994). A delicate balance: Enhancing literature instruction for students of English as a second language. *The Reading Teacher, 47*(6), 438-449.

Goldenberg, C. & Patthey-Chavez, G. (1995). Discourse processes in *Instructional Conversations*: Interactions between teacher and transition readers. *Discourse Processes, 19*, 57-73.

Lee, O., & Fradd, S. H. (1996a). Interactional patterns of linguistically diverse students and teachers: Insights for promoting science learning. *Linguistics and Education: An International Research Journal, 8*, 269-297.

Lopez-Reyna, N. A. (1996). The importance of meaningful contexts in bilingual special education: Moving to whole language. *Learning Disabilities Research & Practice, 11*(2), 120-131.

Minicucci, C., Berman, P., McLaughlin, B., McLeod, B., Nelson, B., & Woodworth, K. (1995). School reform and student diversity. *Phi Delta Kappan, 77(1)*, 77-80.

Perez, B. (1994). Spanish literacy development: A descriptive study of four bilingual whole-language classrooms. *Journal of Reading Behavior, 26*(1), 75-94.

Reyes, M. (1992). Challenging venerable assumptions: Literacy instruction for linguistically different students. *Harvard Educational Review, 62*(4), 427-446.

Ruiz, N. T. (1995). The social construction of ability and disability: II. Optimal and at-risk lessons in a bilingual special education classroom. *Journal of Learning Disabilities, 28*(8), 491-502.

Tuyay, S., Jennings, L., & Dixon, C. (1995). Classroom discourse and opportunities to learn: An ethnographic study of knowledge construction in a bilingual third-grade classroom. *Discourse Processes, 19*, 75-110.

The predominant picture illustrated by this set of 15 studies was that oral language use by English-language learners in the classroom was consistently low. Students had limited opportunities to respond to challenging, higher-order thinking questions, or engage in problem-solving activities that required complex thinking skills. This was equally true when the language of instruction and learning was English, or Spanish, the students' native language in this set of studies. Some of the qualitative studies did note, however, that somewhat more language use occurred when instruction was exclusively in the native language (Lopez-Reyna, 1996; Ruiz, 1995).

Instructional techniques most strongly criticized in descriptive research and observed frequently in these studies include the following practices: (a) asking questions that required 1 or 2 word answers; (b) the exclusive use of whole class instruction with no opportunities for students to work in pairs or cooperative learning small groups (e.g., Arreaga-Mayer & Perdomo-Rivera, 1996; Lopez-Reyna, 1996; Padrón, 1994; Tuyay, Jennings, & Dixon, 1995); and (c) a stress on low cognitive tasks such as copying and on the strictly surface features of language learning, such as capitalization and literal comprehension.

Recommendations in these descriptive studies typically related to increasing student engagement levels or increasing opportunities for student learning. For example, Tikunoff et. al.'s (1991) study of exemplary programs identified that programs considered to be effective tended to utilize high percentages of small group and individual instructional groupings. In fact, Tikunoff et al. did observe high rates of student engagement (96.8%) and task completion behaviors (78.2%).

The set of studies frequently addressed differences in the cultural backgrounds of teachers and students. Terms such as "cultural congruence," "cultural mismatch," "discourse patterns," and "participation styles" were used to describe how aspects of culture enhance and/or hinder student learning. Reyes (1992) and Delpit (1994) are

among the most eloquent in their descriptions of how culture may influence the way students respond to and interpret teacher instructions and instructional styles.

The issue of culturally influenced classroom interactions has not been the subject of experimental research. However, much of the literature describing the academic difficulties of English-language learners contends that culturally influenced interactions are a major factor in how English-language learners perform or fail to perform in classrooms. Recently, this line of interpretive research has expanded to include students with disabilities (Lopez-Reyna, 1996; Ruiz, 1995).

Major Findings from the Professional Work Groups

The professional work groups reinforced that the current state of instructional practice with English-language learners was generally quite poor, and that the knowledge base in this area was extremely limited. Although much of the information provided by these groups was more conjectural than we expected, they did prove to be extremely helpful in sorting through a wide array of issues, themes, and paradoxes raised in the intervention and descriptive studies.

For example, the professional work groups helped us better understand the persistent confusion in research studies between English language development and content acquisition. There was agreement among the groups that Cummins' (1980) distinction between Cognitive Academic Language Proficiency (CALP), now referred to as "academic language" (Cummins, 1984), and conversational skills in English was a useful way of (a) thinking about instruction, (b) actually teaching in the classroom, and (c) communicating with other teachers. There was some consistency among the professional work groups that limited curriculum materials were available for English-language development, especially in the area of merging English-language development with academic content instruction.

We categorized the content of the professional work group "findings" into a category of overarching themes and guidelines, and a category of specific practices that would enhance learning outcomes for English-language learners.

An overarching theme consistently expressed among the groups was that principles of effective teaching for native English-speaking students were effective for English-language learners. The groups also agreed, however, that these principles have to be *modulated* and shaped to meet the simultaneous goals of English language development and content acquisition. In other words, effective teaching with English-language learners was more than just "good teaching," but the general research base on reading and math instruction serves as a solid foundation for effective teaching of English-language learners.

A key to this modulation seems to be that English-language learners needed frequent opportunities to use oral language in the classroom. This daily and active use of language should be structured to include both conversational language and academic discourse. Techniques such as class-wide peer tutoring seem promising.

The groups consistently expressed support for the principle that English-language learners should be taught through the use of challenging material that does not get "watered down" merely because students are not fluent in the language of instruction. However, they frequently commented on how very difficult it is to implement this principle effectively.

Findings from the professional work groups also included the delineation of specific instructional strategies that can be used successfully with English-language learners. Each group (except the first) refined, clarified, and occasionally rejected statements of the earlier groups, and thus the list that emerged reflects to some extent, "group wisdom" about specific aspects of teaching.

Our research team integrated the strategies across the professional work groups, and identified intervention studies and descriptive studies that seemed to illustrate the principles.

These specific instructional strategies and supporting studies are presented in Tables 5 and 6. The principles in Table 5 are grouped together because, in our view, they help elicit aspects of a framework one group referred to as "dynamic, structured teaching."

The principles in Table 6 are grouped together because they represent a major topic in the professional work groups, how to effectively merge English-language development with academic instruction. Many instructional problems have centered on attempts to merge content area learning with English Language Development in so-called "sheltered" approaches. Therefore, we asked each group about specific technologies or principles that they found productive in the integration of language and content knowledge development.

Table 5. Professional Work Group Suggestions Regarding the Importance of an Overall Teaching Framework

Quality Indicator
One key indicator of a quality program is that students are talking, speaking, writing, and reading. These academic behaviors should be observable in every subject area. The following specific points mentioned in the professional work groups further illustrate this feature.
Instructional Principles
1. Utilize teaching structures and formats that elicit frequent student responses and extended student responses. (Echevarria, 1995; Waxman et al., 1994)
2. Utilize more extensive modeling and think-alouds than is found in current practice. (Jiménez, 1997; Muñiz-Swicegood, 1994)
3. Include student and teacher talk, specifically "academic talk," rather than just sharing or conversational talk. Academic talk includes discussion of concepts. (Saunders et al., 1998)

4. Share learning strategies with students: For example, why the teacher has chosen to use a particular learning strategy, labeling the strategy, and telling students why the strategy might be helpful. *Instructional Conversations* can be a technique/strategy for this and for having students talk about important concepts in the content area. (Cardelle-Elawar, 1990; Chamot, Dale, Malley, & Spanos, 1993; Echevarria, 1995; Saunders et al., 1998)

Table 6. Professional Work Group Suggestions on Merging English Language Development with Content-Area Learning

1. English Language Development programs must include development of oral and written proficiency, development of academic language (Cummins, 1994) and basic conversational English, and systematic proactive teaching of conventions and grammar. (Fashola, Drum, Mayer, & Kang, 1996; Saunders et al., 1998; Waxman et al., 1994)
2. In teaching English during English language development, avoid oversimplifying instruction with contrived, intellectually insulting material. In subjects such as science, native language instruction could be confusing to students, since teachers may not have adequate knowledge of technical vocabulary. (Lee & Fradd, 1996b)
3. Use visuals and extensive use of written language to reinforce verbal content when teaching in English. (Rousseau et al., 1993; Saunders et al., 1998)
4. Employ strategic use of synonyms. Teachers' word choice and sentence structure needs to be consistent and concise during second language learning. Teachers also need to pay attention to their use of metaphors and similes and other highly culture-specific phrases and expressions. (Cardelle-Elawar, 1990; Gersten & Jiménez, 1994)
5. Focus on approximately 5-8 core vocabulary words per lesson. Some strategies include: careful selection of words (evocative words to stimulate instruction, key words for understanding a story); linking words or concepts to words known in the native language; showing new words in print; using visuals (e.g., concept maps) to depict concepts or word meanings. (Rousseau et al., 1993; Saunders et al., 1998)
6. Published curriculum materials can serve as effective starting points in promoting English-language learning. However, it is very important to determine if published programs actually provide adequate focus and structure for teaching English. Goldenberg and Sullivan (1995, April) found counter examples with published materials that did not promote English-language learning.
7. During the early phases of language learning, it is important that a teacher modulate and be sensitive in providing feedback and correction on language usage; however, during later stages of language learning, it is important that the teacher identify errors and provide specific feedback to students. (Cardelle-Elawar, 1990)
8. Native language use during English language development must be strategic. At times, it might be useful to use both native language and English during instruction; however, teachers need to be aware of the risk of over-reliance on simultaneous translations. (Klingner & Vaughn, 1996)

Multi-Vocal Integration and Synthesis

At the current time, there is an extremely scant empirical data base to support any specific instructional practices for English-language learners. This is the most cautious and realistic appraisal of the findings from the exploratory meta-analysis.

We stress the lack of empirical support for any specific practice or approach, in part, because current literature in bilingual education and bilingual special education, suggests that much is known about effective instructional practices and strategies for English-language learners. Many studies claim to describe effective practice, yet virtually none of them provide the type of outcome data necessary to draw firm conclusions.

For example, in studies that describe effective practice, both Tikunoff et al. (1991) and Gersten (1996a; 1996b), focused on classrooms that had been "nominated" as exemplary bilingual classrooms. However, there is no empirical evidence that these classrooms enhance student learning, as August and Hakuta (1997) aptly noted. Yet it is equally true that observations of these classrooms using both reliable, low inference measures (Tikunoff et al., 1991) and high inference (qualitative) observations (1996a; Gersten, 1996b; Gersten & Jiménez, 1994), indicated a very different type of bilingual instruction is possible than the instruction described by many researchers (e.g., Ramírez, 1992; Ruiz, 1995). In these exemplary classrooms student language use is relatively high, as is student engagement in learning activities.

The point needs to be emphasized, however, that well-designed and executed experimental studies are needed to uncover the causal links between features of instruction and learning outcomes. Yet during the period from 1985 to 1996, we found a mere eight studies that utilized a valid experimental or quasi-experimental design to explore the impact of instructional strategies on student learning for English-language learners in grades K-8. The fact there were so few studies was corroborated by the recent review by the National Academy of Sciences (August & Hakuta, 1997). Because

there are only a handful of studies, and because the foci among the set of studies varies widely in terms of the age of the participants (kindergarten in Goldenberg et al., 1992 to middle school in Henderson & Landesman, 1995), intervention length (30 days in Cardelle-Elawar, 1990, to one year plus in Henderson & Landesman, 1995, and Saunders et al., 1998) and content focus (e.g., reading, math, language), we could not make any firm generalizations about specific instructional components that lead to enhanced outcomes.

We did note the failure of two widely recommended approaches for teaching English-language learners - *Instructional Conversations* and *Cognitive Academic Language Learning Approach* (i.e., *CALLA*) - to produce the positive outcomes in literacy that one might have anticipated. Both approaches did produce positive effects; *CALLA* produced a small effect on a standardized language measure, and *Instructional Conversations* resulted in a small effect in the structure of story retellings. However, the average effect size in these studies was zero (*CALLA*) or negative (*Instructional Conversations*), and the approaches seemed particularly ineffective in the areas of reading and reading comprehension.

In the case of the *CALLA* in the Waxman et al. (1994) study, part of the problem may have been weak levels of implementation. A large number of teachers in several schools were given limited training in a complex, fundamentally different way of teaching. With *Instructional Conversations* (i.e., Echevarria, 1995), implementation was assessed and proven to be of high quality.

Part of the difficulty with *Instructional Conversations* may have been that the explicit instruction component that advocates of *Instructional Conversations* increasingly advocate (see Saunders & Goldenberg, 1996) may have been missing. Saunders and Goldenberg came to this conclusion after participating with a group of elementary school teachers in a study to describe and implement *Instructional Conversations*. Initially the teachers they worked with had extremely negative views of explicit instruction,

characterizing it as "rigid and formulaic and in direct contrast to the kinds of teaching they wanted to do" (p. 144). The same teachers held very positive views of contructivist teaching practices such as *Instructional Conversations*.

In working with Goldenberg and Saunders (1996), these teachers were asked to develop curriculum and instruction consistent with the state's language arts framework, read and discuss numerous articles on direct instruction and *Instructional Conversations*, and implement *Instructional Conversations* with their English-language learners. Not only did the teachers develop a generally more positive view of explicit instruction over time, but "the relevance of direct teaching emerged consistently as the group worked to identify the elements of *Instructional Conversations* across the year. In fact, direct teaching was written in and detailed as an element of*Instructional Conversations* precisely because the group came to realize that *Instructional Conversations* often require direct teaching" (Saunders & Goldenberg, 1996, p. 152).

For example, one story about a country mouse and city mouse required an understanding of differences between the country and the city, of which many of their students were entirely oblivious. The teachers believed that using *Instructional Conversations* with their English-language learners without first teaching differences between the country and city would result in superficicial discussions and shallow levels of comprehension. As one teacher said "If we don't make sure that they have that deep understanding [of the story], then our later questions could just go over their heads" (p. 151).

This critical aspect of *Instructional Conversations* may have been de-emphasized or completely missing in Echevarria's (1995) study. In any case, we were unable to ascertain that direct teaching was part of the *Instructional Conversations* approach. The outcome, too, is far from conclusive. The extremely small number of students (N=5) and the erratic pattern of results, leading to negative effects in richness of retell and literal comprehension, but positive effects on structure of the retellings, leads us to cautiously

observe that the study overall failed to indicate a clear pattern of either positive or negative effects.

Three Possible Themes of Effective Instruction

Our analysis of the published studies and our work with the professional work groups led to the emergence of three themes of effective instruction for English-language learners. We explore these themes in the remainder of the discussion section.

Approaches that Extend the Knowledge Base on Effective Teaching to English-language Learners

The first theme is that instructional approaches that extended and modulated practices validated in the instructional research literature of the 1980s and earlier 1990s seemed to produce positive effects. These more traditional approaches ranged from *class-wide peer tutoring* (Delaquadri, Greenwood, Whorton, Carta, & Hall, 1986) to Jane Stalling's (1980) system for increasing academic engagement through the increased use of active teaching behaviors (Waxman et al., 1994), to emphasizing accurate decoding and word recognition during reading (Goldenberg et al., 1992). The approaches tend to increase the amount of active engagement in academic learning and/or the quality and quantity of feedback provided to students during lessons.

In our analysis, then, most of the teaching techniques or structures that seemed to lead to enhanced learning outcomes were extensions of the "effective teaching" and peer-mediated instructional research base. These would also include the "tailoring of feedback" used by Cardelle-Elawar (1990) for mathematical problem solving and the provision of focused, explicit instruction on math concepts in Henderson and Landesman (Henderson & Landesman, 1995).

Although techniques such as *Instructional Conversations* and *CALLA* did not produce the academic learning outcomes expected, these studies are clearly attempting to address an important problem articulated in the professional work groups and the descriptive studies–lack of student speech during instruction. Yet finding the link

between increasing student language use and increasing student content area learning remains elusive, as the intervention studies document. We need to know a good deal more about which components of instruction in the intervention studies really contributed to student growth. This problem is most obviously represented in the complex, multi-faceted interventions, such as Saunders et al. (1998).

Therefore, one issue or theme that begins to emerge from the exploratory meta-analysis, and that was reinforced in some of the professional work group discussions, was that the empirical research on instruction for native-English speakers provides the beginnings of a foundation for improving instruction for English-language learners. However, we need to know much more about how to "tailor" or modulate techniques to better "fit" English-language learners.

In the area of literacy, the most interesting overall finding was that methods that increase academic engagement, such as the Stallings method of professional development (i.e., Effective Use of Time, used inWaxman et al., 1994), with its focus on specific teaching techniques to increase academic engagement, tend to increase achievement. This finding was true for both first language literacy (Echevarria, 1995) and second-language literacy (Waxman et al., 1994). The analysis also supports the use of certain specific techniques such as preteaching of critical vocabulary prior to student reading (Rousseau et al., 1993), building background knowledge (Saunders et al., 1998), and providing explicit instruction and guided practice in math problem solving (Cardelle-Elawar, 1990).

In general, abhorrence for more explicit methods of teaching, such as those used in Stallings' *Effective Use of Time* (i.e., Waxman et al., 1994) or Rousseau et al. (1993) can be found in many of the qualitative observational studies, of Ruiz (1995), Lopez-Reyna (1996), and Perez (1994), all of which are of the interpretivist mode (Ferguson, Ferguson, & Taylor, 1992) of qualitative research. Yet these explicit methods seem to be effective if teachers use strategies based on research.

Insights gained from professional work groups. The sequential and focused nature of the professional work groups allowed us to explore predominantly semantic conflicts by consistently working to operationalize members' comments about instruction and seek clarification about what a given approach would look like when implemented. For example, an early professional work group advocated "structured dynamic teaching" as an optimal approach. It was unclear to us, and other members of the group, what this really meant. After a lengthy group discussion, there was consensus that it represented a set of instructional activities during which students have an opportunity to participate in fairly lengthy, complex verbal exchanges with their teachers and peers, and where teacher guidance is clear.

With subsequent groups we presented the term "structured dynamic teaching" and our understanding of its meaning, to that point, and asked the group for their reactions and feedback. In this way we tried to continuously move toward a degree of consensus with the groups about what the important principles and strategies were that constituted effective instruction.

The professional work groups consistently supported a method of instruction we came to label a "hybrid" model that we believe: (a) captures the essence of "structured dynamic teaching," (b) reflects extensions of validated instructional approaches of the Brophy-Good framework, and (c) incorporates principles of teaching emanating from advances in cognitive psychology. One researcher succinctly noted the group's tendency to want to synthesize: "Taking the 'best' of both direct instruction and communicatively-based classroom interaction seems to be the most powerful vehicle towards accomplishing effective and optimal instruction." The critical goal of this approach is the simultaneous development of language and academic proficiency.

Principles of Best Practice

We identified five specific instructional variables or principles from our multi-vocal analysis that, although supported by limited experimental evidence, suggest

critical components for instruction: (a) vocabulary as a curricular anchor, (b) visuals to reinforce concepts and vocabulary (c) cooperative learning and peer tutoring strategies, (d) strategic use of the native language, and (e) modulation of cognitive and language demands. We briefly describe each of these components in this section.

Building and using vocabulary as a curricular anchor. A clear area of consensus among the professional work groups was that vocabulary learning should play a major role in successful programs for English-language learners. The number of new vocabulary terms introduced at any one time should be limited. The standard method of presenting up to 20 or more new vocabulary words that students are expected to learn at a given time is not an effective way to help English-language learners develop vocabulary. Teachers in our professional work groups recommended using lists of 7 or fewer words that students would work on over relatively long periods of time. Criteria for selecting words should be considered carefully, so that words are selected that convey key concepts, are of high utility, that are relevant to the bulk of the content being learned, and that have meaning in the lives of students.

Restricting the number of words students are expected to learn will help them learn word meanings at a deep level of understanding, an important principle of sustained vocabulary growth. The research of Nagy (1988) and Beck and McKeown (1985) were cited as important resources for helping teachers understand how to teach vocabulary to English-language learners. The professional work groups felt that many teachers needed guidance in selecting vocabulary words for instruction, as districts and conventional texts rarely provide the type of guidance needed.

One professional work group member provided insights into the methods she used to select and teach vocabulary that were strongly supported by other members of the group. She noted how she chose words for the class to analyze in depth that represented complex ideas - adjectives like *anxious, generous,* and *suspicious,* and nouns like *memory* - words that English-language learners are likely to need help with and

words that were linked to the story in meaningful and rich ways. Students had to read the story and look for evidence that certain events or descriptions that were connected to vocabulary instruction pertained to a particular character or incident.

Two of the intervention studies had components that dealt specifically with vocabulary development. Vocabulary learning was the explicit focus of the study by Rousseau et al. (1993). Teachers used a variety of methods to teach word meanings to students including visually presenting the words, defining them, and using gestures and other visual techniques (e.g., pictures). It is interesting that both of the outcome measures (i.e., accurate reading of all the words in the story and comprehension of the story) showed dramatic improvement over a method in which teachers previewed the entire story with students by reading it to them.

In Saunders et al. (1998) as well, critical vocabulary were identified prior to story reading. A range of approaches were used to help students develop a deep understanding of these words. Students were also guided to link critical vocabulary to relevant experiences in their lives.

In both studies, the time-tested practice of introducing new vocabulary prior to reading a new story was used successfully. Echevarria (1998) described how this type of vocabulary instruction can be used with English-language learners: "One form of vocabulary development includes short, explicit segments of a class time in which the teacher directly teaches key vocabulary. These five minute segments would consist of the teacher saying the vocabulary word, writing it on the board, asking students to say it and write it and defining the term with pictures, demonstrations, and examples familiar to students" (p. 220).

Use of visuals to reinforce concepts and vocabulary. Two of the professional work group discussions focused particularly on the importance of using visuals during instruction. These visuals might range from complex semantic visuals (Reyes & Bos, 1998), to visuals based on text structures, such as story maps and compare-contrast

"think sheets." Visuals are especially successful in supporting English language development because they are such a good way to help students visualize the abstractions of language.

Two of the intervention studies and several of the observational studies noted that the use of visuals during instruction increased learning. Rousseau et al. (1993) used visuals for teaching vocabulary (i.e., words written on the board and the use of pictures), and Saunders et al. (Saunders et al., 1998) incorporated the systematic use of visuals for teaching reading and language arts. Visuals also typically play a large role in *Cognitive Academic Language Learning Approach* [*CALLA*], although Waxman et al. (1994) did not specifically indicate how visual organizers were used in their study.

The double demands of learning *content* and a *second-language* are significant; the difficulty should not be underestimated. Because the spoken word is fleeting, visual aides such as graphic organizers, concept and story maps, and word banks give students a concrete system to process, reflect on, and integrate information.

Implementation of even simple techniques such as writing key words on the board or a flip chart while discussing them verbally can support meaningful English language development and comprehension. The professional work groups concurred that even the simple integration of visuals is drastically under-utilized, and even when used, methods are typically inconsistent or superficial and do not support students' deep processing and thinking.

Further research on how to use visuals to enhance English-language learning is needed. Also, because of the consistent, strong support for the use of visuals expressed in the professional work groups, we believe educators involved in professional and curriculum development, or curriculum selection, should seriously consider this issue, as well.

Use of cooperative learning and peer tutoring strategies. We believe cooperative learning and peer tutoring strategies have the potential to effectively and rapidly

increase English-language development, particularly decontextualized language concepts with high degrees of cognitive challenge. One of our original propositions was that certain *specific techniques* in cooperative learning lead to superior student outcomes. In the professional work groups, the need for highly structured cooperative learning groups was often stressed.

Two of the intervention studies used cooperative learning or peer tutoring strategies as critical pieces of their interventions. Klingner and Vaughn (1996) tested whether cooperative learning or peer tutoring was more effective in promoting comprehension with English-language learners with learning disabilities. Although there was some evidence that peer tutoring was the most effective, both of the interventions led to improved learning outcomes. In the intervention used by Muñiz-Swicegood (1994), students worked in successively smaller cooperative groups (until they were finally working in pairs) to learn how to generate and answer questions about what they were reading. Students in this intervention did better on measures of reading comprehension than students who were taught using basal reading approaches.

Strategic use of the native language. Strategic use of students' native language can help insure that the development of higher-order thinking skills receives adequate curriculum focus. The professional work groups agreed with the general concept that a viable way to achieve this objective is for teachers to use levels of English that students are very fluent with, while simultaneously using more extensive native language to introduce complex concepts and provide students with opportunities to concentrate on understanding challenging context. The professional work groups, however, failed to reach consensus on *how* students' use of their native language could be used strategically for this purpose. This issue was discussed in many of the descriptive studies reviewed (e.g., Gersten, 1996b; Lopez-Reyna, 1996; Ramírez, 1992; Tikunoff et al., 1991).

The strategic use of native language is a controversial issue. Most researchers from the professional work groups cautioned against using dual translations frequently, that is, the extensive use of both the student's native language and second language during instruction. However, one researcher advocated a counter position, suggesting that written words be provided in both English and Spanish. Many researchers from the observational studies (Gersten & Jiménez, 1994; Lopez-Reyna, 1996; Miniccuci et al., 1995; Tikunoff et al., 1991) proposed using a student's native language as an instructional approach. Yet, the observational findings of Ramírez (1992) indicate that neither more nor less higher-order discussion occurred when instruction was in the native language or in English. Thus, our conclusion is that it is beneficial to use students' native language, but it should be done in a strategic fashion, and in general avoid the tendency to provide dual translations.

Two of the intervention studies incorporated the strategic use of native language to help with learning difficulties in the second-language. In the Cardelle-Elawar (1990) study, very focused attention was devoted to exploring the meaning of the language used in the math story problems and how students could use a variety of strategies, including their knowledge of Spanish, to help them understand and figure out what the problem in English was asking them to do. This type of intense instruction to determine what specifically is being requested in a problem solving situation led to very large effects compared to broader instructional approaches. In Klingner and Vaughn (1996), students were encouraged to use their native language strategically to solve specific problems they were encountering in their cooperative learning and peer tutoring groups.

Modulation of cognitive and language demands. This last instructional strategy carries a different weight of importance, and we view it as the most speculative among those we have proposed. One of the propositions shared with each of the professional work groups was that during English language content instruction, effective teachers

intentionally vary the cognitive and language demands. Typically, there is an inverse relationship between the two. When cognitive demands are high, language expectations are simplified, and teachers, for example, may accept brief or truncated responses in English. In another part of the lesson, cognitive demands are intentionally reduced so that students can more comfortably experiment with extended English language use.

This proposition was supported in each of the five professional work groups and appears consonant with contemporary theories of second-language acquisition (e.g., August & Hakuta, 1997). Obviously, empirical support for this proposition is needed, although designing a suitable research study around such a subtle principle will be difficult.

Confusion, Tension, and Assumptions about Oral Language Use

The second major theme that emerged from the multi-vocal synthesis involves concerns and confusions about the role of oral language in academic instruction. All the studies describing classroom learning envirnoments (both low and high inference) noted rare student oral use in the classroom. This issue was stressed both in studies of English language development and in studies of native language content instruction.

In the following section, we argue that both extended discourse about academic topics and briefer responses to specific questions about content are cornerstones of academic growth for English-language learners. We believe this is a valid interpretation, based on trends in the research studies and our interactions with the professional work groups.

Our review of the data sources suggests that discussions of potentially effective instructional practices for English-language learners over-emphasize natural language use and do not clearly articulate the important distinctions involved when language use is the major goal and when cognitive or academic growth is paramount. To understand this confusion, we review some of the observational research.

Relevant findings from research. Ramírez (1992) described typical classrooms as passive learning environments for students. Teachers do the majority of talking and student contributions are in response to teacher questions. Other studies support this pattern (e.g., Arreaga-Mayer & Perdomo-Rivera, 1996; Lopez-Reyna, 1996; Padrón, 1994).

Specifically, Ramírez (1992) reported that student language use and opportunities to engage in cognitively challenging tasks were extremely low. In his observations, the mean proportion of student-initiated language use ranged from .3 to 10.1 percent of the total time in which students were responding. This low rate of student-initiated responses was corroborated in the high-inference, qualitative observational studies reviewed (Lopez-Reyna, 1996; Perez, 1994; Ruiz, 1995), where student discourse was typically limited to one or two word utterances.

Perhaps most astounding is the low level of student oral language use in English Language Development classes noted by Arreaga-Mayer and Perdomo-Rivera (1996). They found that only 21% of the time did observed students use written or oral language. In other words, students rarely spoke during classes in which the explicit purpose was English Language Development. A tension that emerges from the literature is the implicit assumption of most researchers that increased language use (be it in the students' native language or in English) should be a high priority goal because it will lead to increased learning. For example, as a rationale for *Instructional Conversations*, Echevarria (1995) wrote that "language is a primary vehicle for intellectual development" (p. 537), and implicit in the philosophy of *Instructional Conversations* is the assumption that increased oral language use by students during reading instruction will improve comprehension.

Yet, *Instructional Conversations* produced negative results on two crucial measures of reading comprehension when contrasted with the type of instruction typical in a more traditional basal reading lesson. In trying to account for the findings,

Echevarria (1995) observed that "While it was speculated that the enriched language opportunities that *Instructional Conversations* provide would enhance the students' narrative construction [i.e., elements of story grammar], it is possible that what takes place in the classroom does not contribute to narrative development [i.e., richness of idea units]." She continued by noting that "the discourse rules of the basal treatments tended to elicit more *who, what , where* types of questions . . . while the *Instructional Conversations* discourse attempted to evoke opinion and more complex language . . . " (p. 550). An interpretation of the results is that the basal intervention tended to create more opportunities for analytic discussions.

In other words, increased use of sophisticated language constructions in school may or may not be related to increased academic and cognitive growth. We simply do not have an empirical knowledge base to inform us as to which of the following forms of student engagement provides greater overall benefit for English-language learners: (a) generous opportunities for oral language interactions, (b) reading, (c) writing, (d) listening, (e) content area activities such as those involved in math or science, or (f) the optimal combination of any of the above.

In the professional work groups, we noted that members often seemed confused by - or vacillated between - two objectives: (a) language learning, in either the native or second-language, and (b) content area learning. We do not wish to imply that oral language use in school is an unimportant objective, or that increased use of oral language is inversely related to academic growth in content areas. Rather, we emphasize that these are two distinct goals, and researchers and educators need to be clear about the distinction. Furthermore, findings in some of the descriptive research (Jiménez & Gersten, in press; Lee & Fradd, 1996a; Ruiz, 1995) indicate that increased student dialogue in class can lead to discussions with minimal cognitive challenge and minimal academic content.

Problems in implementation of the intervention approaches that require extensive natural language and authentic dialogue, such as *CALLA* and *Instructional Conversations*, may help explain why they failed to lead to effects in reading. The implementation problems that plagued the large scale research study by Waxman et al. (1994), for example, were consistently corroborated in the professional work groups. Participants talked about weak, inconsistent, and sometimes incoherent implementation of techniques such as semantic mapping, cooperative learning, and story mapping. One member of the California work group noted that techniques such as semantic mapping and teachers' thinking aloud "all have been used non-effectively in recent years." Extensive discussion in three of the professional work groups addressed weak implementation of cognitively-based approaches and limited curricula or manuals available for teacher use.

Likewise, in response to complaints about weak implementation of cooperative learning, one teacher-researcher indicated that by using highly structured groups, she virtually never had the kinds of problems that others discussed as chronic and endemic. In other words, by using established principles of effective instruction such as clear expectations, frequent monitoring, and immediate feedback to students, this teacher was able to overcome seemingly intractable problems in using an innovative practice to increase language use. We believe one reason the highly structured *Classwide Peer Tutoring* [CWPT] method (Klingner & Vaughn, 1996) surpassed the more loosely structured *Cooperative Learning* method was because student roles and task demands were more clearly explained and monitored in the former than in the latter.

Merging English Language Development with Content-Area Learning

The third major theme of the multi-vocal synthesis that emerged primarily in the professional work groups was encouragement for the increased use of approaches such as *sheltered content area instruction*. We think there were many reasons for this, an important one being that in some districts there are so many language groups that

native language instruction for all or even a majority of English-language learners is not always feasible. A second reason was the trend towards providing students with specific language assistance in grades 3-6, a time when they leave classes that provide predominantly native language instruction and move into classrooms that are conducted primarily in English.

We want to clarify that we are not advocating for the exclusive use of instructional approaches that merge English Language Development with content-area learning (i.e., in opposition to strategic use of native language instruction). Rather, we advocate for the strategic use of these approaches and hope to uncover some of the current problems, as well as identify specific strategies teachers can use to promote English Language Development during academic instruction. Because of the relative novelty of this approach in American schools, discussions in the professional work groups on this topic were often very rich.

We invested a good deal of energy in trying to understand the histories of the various approaches to English language development (e.g., formal/syntactic, natural language, and sheltered content area), in part because during the first two professional work groups, participants admitted that "definitions of English-as-a-Second-Language and sheltered instruction are unclear."

The rationale for Content Area ESOL instruction is that students can learn English while learning academic content, and that this type of learning will build academic language (Cummins, 1994) because students will be learning the abstract language of scientific or mathematical or literary discourse. Furthermore, Content Area ESOL is better suited to classrooms than natural language since classrooms are, by and large, places of learning.

However, the professional work groups were consistent in indicating that:

- content area instruction often leads to sacrifices in learning English;

- few districts have a curriculum program or approach that promotes students' proper use of the English language.

In discussing concerns about instruction for English-language learners, professional work group members frequently noted how Content Area ESOL almost invariably fails, in the words of one group member, to provide "adequate time for English language learning." In other words, participants felt that teachers often emphasize content acquisition over building English language abilities.

In fact "the dilution of ESOL instruction" under the term sheltered content area instruction, and the overall neglect of ESL/ESOL instruction, was a recurrent refrain in the California professional work group. One teacher noted "It's important to use content as a basis for language development . . . [however] there is a risk during content instruction of neglecting language development" (California professional work group, October, 1996). Another educator from the district bilingual education office noted that "It's important for teachers to be clear about objectives and goals . . . yet an explicit statement of goals does not exist [in district or state curricula materials]." Some members suggested a set of curriculum goals that include "specific language concepts," noting how so many teachers merely "hope that language occurs" during content area lessons.

One researcher in the group noted how the need for explicit teaching in ESL/English Language Development classes "should never be underestimated . . . " He stressed "the importance of promoting language while promoting thought," voicing the concern that students need experiences in "thinking through" and then verbalizing their ideas in content areas (e.g., science, mathematics, history), in English. Attempts to merge content area instruction with ESOL instruction, though well intended and conceptually sound, are rarely well implemented.

The major problem highlighted in Content Area ESOL was how time for language learning often is truncated or omitted. These concerns are reflected in the

data from observational studies by Ramírez (1992) and Arreaga-Mayer and Perdomo-Rivera (1996). Arreaga-Mayer and Perdomo-Rivera (1996) both noted how the general education and the ESL settings failed to provide instruction to facilitate second-language acquisition. Similarly, Ramírez (1992) concluded that in all the varying models of bilingual education, teachers did not promote language development effectively. He stated that "consistently, across grade levels within and between the three instructional programs, students are limited in their opportunities to produce language and in their opportunities to produce more complex language" (p. 9).

This pattern also supports a major finding in our study of issues confronting teachers in the upper elementary grades (Gersten, 1996a, 1996b), and also found in observational research by Reyes (1992). We see inadequate time for English language development as a major problem with current practice.

Several reasons for this problem were identified in the professional work groups. First and foremost was teachers' concern for increased accountability for content learning (as measured by test results), as opposed to the more amorphous goals of English language acquisition, and a relative de-emphasis in accountability for students' language development needs. Participants in the professional work groups discussed in detail how, based on their observations and experiences in classrooms, the tendency to cover all the content in science or social studies or mathematics almost invariably precluded allowing adequate time for English Language Development, especially more formal academic English.

Other comments in the professional work groups focused on failure to systematically impart to students skills in speaking and writing standard English, even in middle school. While many members felt that the policy of never correcting students for grammatical or pronunciation problems during English-language instruction made sense during the early years of English Language Development, there was general consensus that students need feedback on their formal English usage as they progress in school, and teachers lack any kind of coherent system for providing it. One

professional work group suggested that in the early phases of language learning, teachers should modulate the feedback they provide students, and be sensitive to the problems inherent in correcting every grammar mistake students make; however, during later stages, one member reflected the feeling in the group by noting the "importance of identifying errors and providing specific feedback."

A recent research study by Fashola, Drum, Mayer, and Kang (1996) may provide some direction in this area. Fashola and colleagues (1996) noted how errors made by Latino students in English are usually predictable, and how these predictable errors could become the basis of *proactive curricula:* "Rather than simply marking a predicted error as incorrect, the teacher could explicitly point out that the phonological or orthographic rule in English is different from the one in "Spanish" (Fashola et al., 1996, p. 840). After reviewing these issues with professional work groups, and reading about problems with Content Area ESOL in sources as diverse as the *New York Times* and the *Harvard Educational Review* (Reyes, 1992), we concluded that an effective English Language Development program should include a component devoted to helping students learn how to use the second-language according to established conventions of grammar and syntax.

We encourage researchers and educators to consider language learning and content-area learning as distinct educational goals, rather than assuming that increased use of oral language in school will automatically lead to an increase in academic learning and the development of higher-order thinking skills. Artful and skillful blending of genuine dialogue, about literature or science, and cognitive challenge is an admirable, but perhaps only occasionally realized goal. On the other hand, providing some time each day when English-language learners have opportunities to work on all aspects of English Language Development, and providing academically challenging content instruction (be it in native language or English), are likely to be more easily achievable, especially if teachers take time to make goals clear and precise.

In short, instruction for English-language learners should work to blend oral language engagement and intellectual (or cognitive) engagement. These distinctions are also important for those doing instructional research in classroom settings. For example, Saunders and colleagues (1998) describe instructional units characterized by high frequency of oral language engagement, but also note that they "view the intellectual substance of the literature units as the driving force in our program" (p. 29).

Summary And Conclusions

1. *We found a total of eight studies* that used valid experimental and quasi-experimental designs to investigate the effects of instructional variables on student learning outcomes with English-language learners. Most of these studies were published in 1994 or later. This lack of an empirical knowledge base should be taken into account when districts or schools are mandated to implement a specific procedure based on "expert knowledge." The knowledge in this area is limited. There are many theories, but very little empirical data.

2. *Within the eight empirical studies,* no clear pattern emerged regarding effective instructional practices with English-language learners. We suggested that there might be a trend supporting instructional approaches that extend effective teaching techniques of the 1980s, (i.e., in the more classic view of the findings by Stallings [1980] and Good and Grouws []).

3. *Studies were often unclear regarding*: (a) how interventions were implemented, (b) the level of implementation that was achieved, (c) the language of instruction, and (d) many other "context" variables that would have given a rich picture to intervention research. We remain convinced that the field must better define interventions, and the critical context variables that give them shape and definition.

4. *Distinguishing between language growth and academic growth* is difficult and needs to be more clearly studied and accounted for. There does seem to be an implicit assumption that suggests that increased language use in the classroom leads to

increased academic growth. The studies did not support this assumption. If anything, there was a small amount of evidence supporting an inverse relationship between language use and academic growth. The issue is a persistent source of confusion in understanding and interpreting studies, and in instructional programs.

5. *The English Language Development* aspect of bilingual education and bilingual special education is cited as a major problem, especially for special education students who may be excluded because they can not keep up with the pace.

6. *We concluded that a good English Language Development program* should include three components. First, one component would focus on the development of proficiency and fluency in English. Both social communication and academic communication of concepts and knowledge that students have previously learned would be addressed. A second component would address the more formal, grammatical aspects of English use. A third component would focus on learning new academic content. In this component, content acquisition would be merged with English acquisition. In contrast to the first component, the content learning demands would be high and the language demands lower. It is important to stress that special education students, many of whom have language related disabilities, need this type of instruction, and should be in programs that include all three aspects. Lack of quality published curricula (as opposed to materials from the military and foreign service) in this area is a major problem.

7. *Regarding future research,* the key is well-designed and valid studies. Federal support has not been strong in this area. Many researchers eschew this population because of the intricacies of measurement. There is no question that there is a limited understanding of the difficulty and complexity of this type of research. The US Department of Education should be made aware of the lack of research and of the difficulties of doing good research in this area.

8. *The work groups with educational professionals* resulted in a set of principles and practices that, we believe, can be very useful in beginning to define best practice. These principles and practices, in particular, highlighted the merger of English Language Development instruction with content area learning, which is increasingly used in American schools. For the most part, these principles were consonant with findings from the empirical exploratory meta-analysis. Most assuredly, they should be part of a research agenda.

Author Notes

This paper was supported by grant number HO23E50013 from the US Department of Education, Office of Special Education Programs. Support was also provided by National Association of State Directors of Special Education. The views expressed, however, represent the authors' and do not necessary reflect the views of the US Department of Education, the National Association of State Directors of Special Education.

We would like to thank the following reviewers who provided insightful and thought-provoking commentary on an earlier version of the manuscript: Robert Rueda, Rose-Marie Weber, Harris Cooper, and Bernice Wong. The authors also would like to thank Ellen Schiller of the US Department of Education for her ongoing support.

We would also like to thank our colleagues Sylvia Smith, Mark Harniss, Janet Otterstedt, and Batya Elbaum, who provided extremely helpful feedback on previous versions of this paper. Finally, we are extremely grateful to the contributions of Susan Marks, who played a large role in many important facets of this synthesis.

References

Arreaga-Mayer C., & Perdomo-Rivera, C. (1996). Ecobehavioral analysis of instruction for at-risk language-minority students. *Elementary School Journal, 96*, 245-258.

August D., & Hakuta, K. (1997). *Improving schooling for language-minority children*. Washington, DC: National Academy Press.

Beck I. L., & McKeown, M. G. (1985). Teaching vocabulary: Making the instruction fit the goal. *Educational Perspectives, 23*(1), 11-15.

Blaunstein P. (1995). The best-kept secrets don't change education: Social marketing presentation. Presentation at OSEP Cross-Project Meeting: Technology Educational Media and Materials for Individuals with Disabilities. November 1-2, 1995, Washington, DC .

Brophy J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Witrock (Ed.), *The third handbook of research on teaching* (pp. 328-375). New York: McMillan.

Cardelle-Elawar M. (1990). Effects of feedback tailored to bilingual students' mathematics needs on verbal problem solving. *Elementary School Journal, 91*(2), 165-175.

Chamot A. U., Dale, M. O., Malley, J. M., & Spanos, G. A. (1993). Learning and problem solving strategies of ESL students. *Bilingual Research Journal, 16*(3 & 4), 1-34.

Chamot A. U., & O'Malley, J. M. (1989). The cognitive academic language learning approach. In P. Rigg & V. Allen (Eds.), *When they don't all speak English* (pp. 108-125). Urbana, IL: National Council of Teachers of English.

Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.

Cummins J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*(2), 175-187.

Cummins J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. San Diego: College-Hill Press.

Cummins J. (1994). Primary languge instruction and the education of languge minority students. In *Schooling and language minority students: A theoretical framework*. Los

Angeles: California State University, National Evaluation, Dissemination and Assessment Center. .

Delaquadri J., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Classwide peer tutoring. *Exceptional Children, 52*, 535-542.

Delpit L. (1994). *Other people's children*. New York: Basic Books.

Echevarria J. (1995). Interactive reading instruction: A comparison of proximal and distal effects of Instructional Conversations. *Exceptional Children, 61*, 536-552.

Echevarria J. (1998). Preparing text and classroom materials for English-language learners: Curriculum adaptations in secondary school settings. In R. Gersten & R. Jiménez (Eds.), *Promoting Learning for Culturally and Linguistically Diverse Students: Classroom Applications from Contemporary Research* (pp. 210-229). Pacific Grove, CA: Brooks/Cole Publishing.

Echevarria J., & Graves, A. (1998). *Sheltered content instruction: Teaching English-language learners with diverse abilities*. Des Moines, IA: Allyn & Bacon.

Echevarria J., & McDonough, R. (1993). *Instructional conversations in special education settings: Issues and accommodations* . Washington, DC: National Center for Research on Cultural Diversity and Second Language Learning.

Fashola O. S., Drum, P. A., Mayer, R. E., & Kang, S. (1996). A cognitive theory of orthographic transitions: Predictable errors in how Spanish-speaking children spell English words. *American Educational Research Journal, 33*, 825-844.

Ferguson P. M., Ferguson, D. L., & Taylor, S. J. (1992). *Interpreting disability: A qualitative reader*. New York: Teachers College Press.

Figueroa R. A., Fradd, S. H., & Correa, V. I. (1989). Bilingual special education and this special issue. *Exceptional Children, 56*(2), 174 - 178.

Fitzgerald J. (1995). English-as-a-second-language learner's cognitive reading: A review of research in the United States. *Review of Educational Research, 65*(2), 145-190.

Gersten R. (1996a). The double demands of teaching English language learners. *Educational Leadership, 53*(5), 18-22.

Gersten R. (1996b). Literacy instruction for language-minority students: The transition years. *Elementary School Journal, 96*(3), 227-244.

Gersten R., & Jiménez, R. (1994). A delicate balance: Enhancing literacy instruction for students of English as a second language. *The Reading Teacher, 47*(6), 438-449.

Gersten R., Morvant, M., & Brengelman, S. (1995). Close to the classroom is close to the bone: Coaching as a means to translate research into classroom practice. *Exceptional Children, 62*(1), 52-66.

Gersten R., & Woodward, J. (1994). *Lost opportunities: Observations of the education of language minority students.* Paper presented at the annual conference of American Educational Research Association, New Orleans, LA.

Gersten R., & Woodward, J. (1995). A longitudinal study of transitional and immersion bilingual education programs in one district. *Elementary School Journal, 95*(223-240).

Glass G., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Goldenberg C., & Patthey-Chavez, G. (1995). Discourse processes in Instructional Conversations: Interactions between teacher and transition readers. *Discourse Processes, 19*, 57-73.

Goldenberg C., Reese, L., & Gallimore, R. (1992). Effects of literacy materials from school on Latino children's home experiences and early reading achievement. *American Journal of Education, 100*(4), 497-536.

Goldenberg C., & Sullivan, A. (1995, April). *Making change happen in a language-minority school: A search for coherence.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Good T. L., & Grouws, D. A. (1977). Teaching effects: A process product study in fourth grade mathematics classrooms. *Journal of Teacher Education, 28*(3), 49-54.

Henderson R. W., & Landesman, E. W. (1995). Effects of thematically integrated mathematics instruction on students of Mexican descent. *Journal of Educational Research, 88*(5), 290.

Hiebert J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*, 393-425.

Jiménez R. T. (1997). The strategic reading abilities and potential of five low-literacy Latina/o readers in middle school. *Reading Research Quarterly, 32*(3), 224-243.

Kelly B., Gersten, R., & Carnine, D. (1990). Student error patterns as a function of curriculum design. *Journal of Learning Disabilities, 23*(1), 23-32.

Klingner J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *Elementary School Journal, 96*, 275-293.

Lee O., & Fradd, S. H. (1996a). Interactional patterns of linguistically diverse students and teachers: Insights for promoting science learning. *Linguistics and Education: An International Research Journal, 8*, 269-297.

Lee O., & Fradd, S. H. (1996b). Literacy skills in science performance among culturally and linguistically diverse students. *Journal of Research in Science Teaching, 32*, 797-816.

Lopez-Reyna N. A. (1996). The importance of meaningful contexts in bilingual special education: Moving to whole language. *Learning Disabilities Research & Practice, 11*(2), 120-131.

Lucas T., Henze, R., & Donato, R. (1990). Promoting the success of Latino language-minority students: An exploratory study of six high schools. *Harvard Educational Review, 60*(315-340).

Mastropieri M. A., Scruggs, T. E., Bakken, J. P., & Whedon, C. (1996). Reading comprehension: A synthesis of research in learning disabilities. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* . Greenwich, CT: JAI Press.

Miles M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills, CA: Sage Publications.

Miniccuci C., Berman, P., McLaughlin, B., McLeod, B., Nelson, B., & Woodworth, K. (1995). School reform and student diversity. *Phi Delta Kappan, 77*(1), 77-80.

Moll L. C. (1988). Some key issues in teaching Latino students. *Language Arts, 65*(5), 465-472.

Muñiz-Swicegood M. (1994). The effects of megacognitive reading strategy training on the reading performance and student reading analysis strategies of third grade bilingual students. *Bilingual Research Journal, 18*(1-2), 83-97.

Nagy W. E. (1988). *Teaching vocabulary to improve reading comprehension*. Neward, DE: International Reading Association.

Natriello G., McDill, E. L., & Pallas, A. M. (1990). *Schooling disadvantaged children: Racing against catastrophe*. New York: Teachers College Press.

Noblit G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies. Qualitative research methods series 11*. Newbury Park: Sage.

Ogawa R. T., & Malen, B. (1991). Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. *Review of Educational Research, 61*, 265-286.

Padrón Y. N. (1994). Comparing reading instruction in Hispanic/limited-English-proficient schools and other inner-city schools. *Bilingual Research Journal, 18*(1), 49-66.

Patton M. Q. (1990). *Qualitative evaluation and research methods*. (2nd ed.). London: Sage.

Perez B. (1994). Spanish literacy development: A descriptive study of four bilingual whole-language classrooms. *Journal of Reading Behavior, 26*(1), 75-94.

Pressley M. (1996). Embracing complexity: How and what I have learned about outstanding literacy instruction in the elementary grades, *Presentation at the 1996 American Educational Research Association annual meeting* . New York.

Ramírez J. D. (1992). Executive summary: Longitudinal study of structured English immersion strategy, early-exit and late-exit transitional bilingual education programs for language-minority children. *Bilingual Research Journal, 16*(1), 1-62.

Reyes E., & Bos, C. (1998). Interactive semantic mapping and charting: enhancing content area learning for language minority students. In R. G. R. Jiménez (Ed.), *Promoting learning for culturally and linguistically diverse students: Classroom applications from contemporary research* (pp. 133-152). Belmont, CA: Wadsworth.

Reyes M. (1992). Challenging verable assumptions: Literacy instruction for linguistically different students. *Harvard Educational Review, 62*(4), 427-446.

Rosenshine B. (1986). Synthesis of research on explicit teaching. *Education Leadership, 12*, 85-92.

Rosenshine B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*(4), 479-530.

Rosenthal R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook for research analysis* . New York: Russell Sage Foundation.

Rousseau M. K., Tam, B. K. Y., & Ramnarain, R. (1993). Increasing reading proficiency of language-minority students with speech and language impairments. *Education and Treatments of Children, 16*, 254-271.

Ruiz N. T. (1995). The social construction of ability and disability: II. Optimal and at-risk lessons in a bilingual special education classroom. *Journal of Learning Disabilities, 28*(8), 491-502.

Saunders W., & Goldenberg, C. (1996). Four primary teachers work to define constructivism and teacher-directed learning: Implications for teacher assessment. *Elementary School Journal, 97*(2), 139-161.

Saunders W., O'Brien, G., Lennon, D., & McLean, J. (1998). Making the transition to English literacy successful: Effective strategies for studying literature with transition students. In R. Gersten & R. Jimenez (Eds.), *Effective strategies for teaching language minority students* (pp. 99-132). Belmont, CA: Wadsworth.

Stallings J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Leadership, 9*(11), 11-16.

Strauss A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* . Thousand Oaks: Sage.

Swanson H. L. (in press). What instruction works for students with learning disabilities? Results from a meta-analysis of intervention studies. In R. Gersten, E. Schiller, S. Vaughn, & J. Schumm (Eds.), *Issues and research in special education* . Mahwah, NJ: Erlbaum.

Tharp R. G., & Gallimore, R. (1988). *Rousing minds to life*. Cambridge, UK: Cambridge University Press.

Tikunoff W. J., Ward, B. A., van Broekhuizen, L. D., Romero, M., Castaneda, L. V., Lucas, T., & A., K. (1991). *Final Report: A descriptive study of significant features of exemplary special alternative instructional programs* . Los Alamitos, CA: The Southwest Regional Educational Laboratory.

Tuyay S., Jennings, L., & Dixon, C. (1995). Classroom discourse and opportunities to learn: An ethnographic study of knowledge construction in a bilingual third-grade classroom. *Discourse Processes, 19*, 75-110.

Vaughn S., Schumm, J. S., & Sinagub, J. M. (1996). *The focus group interview: Use and application in educational and psychological research*. Newbury Park, CA: Sage.

Waxman H. C., de Felix, J. W., Martinez, A., Knight, S. L., & Padrón, Y. (1994). Effects of implementing classroom instructional models on English language learners' cognitive and affective outcomes. *Bilingual Research Journal, 18*(3-4), 1-22.

Wolcott H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks: Sage.

Wortman P. M. (1994). Judging research quality. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* . New York: Russell Sage Foundation.

Wortman P. M., & Bryant, F. B. (1985). School desegregation and black achievement: An integrative review. *Sociological Methods and Research, 13*, 289-324.

Yzaguirre R. (1998,). What's wrong with bilingual education?: Is it 'lingual or is it 'education'? *Education Week*, pp. 72, 46-47.

---

[1] The fact there were only 8 studies in the intervention category was recently confirmed in a report by the National Academy of Sciences (August & Hakuta, 1997). Because of the small number of studies, we consider this an *exploratory meta-analysis*. One single-subject study (Rousseau, Yung Tam, & Ramnarain, 1993) met our over-arching purpose for the quantitative analysis (i.e., to find empirical support for beneficial instructional practices), but in not using a between-groups design was not amenable to meta-analysis techniques. Thus, we included this study in our pool of quantitative studies, but analyzed it separately from the exploratory meta-analysis.